# Matrix Diagonalization

Miliyon T.

August 25, 2016

## 1 Introductions

When we introduced eigenvalues and eigenvectors, we wondered when a square matrix is similarly equivalent to a diagonal matrix? In other words, given a square matrix $A$, does a diagonal matrix $D$ exist such that $A \sim D$? (i.e. there exists an invertible matrix $P$ such that $A = P^{-1}DP$)

In general, some matrices are not similar to diagonal matrices. For example, consider the matrix

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

Assume there exists a diagonal matrix $D$ such that $A = P^{-1}DP$. Then we have

$$A - \lambda I_n = P^{-1}DP - \lambda I_n = P^{-1}DP - \lambda P^{-1}P = P^{-1}\Big(D - \lambda I_n\Big)P,$$

i.e $A - \lambda I_n$ is similar to $D - \lambda I_n$. So they have the same characteristic equation. Hence $A$ and $D$ have the same eigenvalues. Since the eigenvalues of $D$ of the numbers on the diagonal, and the only eigenvalue of $A$ is 2, then we must have

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2I_2.$$

In this case, we must have $A = P^{-1}DP = 2I_2$, which is not the case. Therefore, $A$ is not similar to a diagonal matrix.

## 2 Diagonalization

**Definition 1.** A matrix is diagonalizable if it is similar to a diagonal matrix.

**Remark 2.1.** The matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

has three different eigenvalues. We also showed that A is diagonalizable. In fact, there is a general result along these lines.

**Theorem 2.2.** Let $A$ be a square matrix of order $n$. Assume that $A$ has $n$ distinct eigenvalues. Then $A$ is diagonalizable. Moreover, if $P$ is the matrix with the columns $C_1, C_2, \cdots, C_n$ the $n$ eigenvectors of $A$, then the matrix $P^{-1}AP$ is a diagonal matrix. In other words, the matrix $A$ is diagonalizable.

Problem: What happened to square matrices of order $n$ with less than $n$ eigenvalues?
We have a partial answer to this problem.

**Theorem 2.3.** Let $A$ be a square matrix of order $n$. In order to find out whether $A$ is diagonalizable, we do the following steps:

1. Write down the characteristic polynomial

$$p(\lambda) = \det\Big(A - \lambda I_n\Big).$$

2. Factorize $p(\lambda)$. In this step, we should be able to get

$$p(\lambda) = (\lambda - \lambda_1)^{n_1} \cdot (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$$

   where the $\lambda_i$, $i = 1, \cdots, k$, may be real or complex. For every $i$, the powers $n_i$ is called the (**algebraic**) multiplicity of the eigenvalue $\lambda_i$.

3. For every eigenvalue, find the associated eigenvectors. For example, for the eigenvalue $\lambda_i$, the eigenvectors are given by the linear system

$$A \cdot X = \lambda_i X \text{ or } \Big(A - \lambda_i I_n\Big)X = \mathcal{O}.$$

   Then solve it. We should find the unknown vector $X$ as a linear combination of vectors, i.e.

$$X = \alpha_1 C_1 + \alpha_2 C_2 + \cdots + \alpha_{m_i} C_{m_i}$$

   where $\alpha_j$, $j = 1, \cdots, m_i$ are arbitrary numbers. The integer $m_i$ is called the **geometric** multiplicity of $\lambda_i$.

4. If for every eigenvalue the algebraic multiplicity is equal to the geometric multiplicity, then we have

$$m_1 + m_2 + \cdots + m_k = n$$

   which implies that if we put the eigenvectors $C_j$, we obtained in 3. for all the eigenvalues, we get exactly $n$ vectors. Set $P$ to be the square matrix of order $n$ for which the column vectors are the eigenvectors $C_j$. Then $P$ is invertible and

$$P^{-1} \cdot A \cdot P$$

   is a diagonal matrix with diagonal entries equal to the eigenvalues of $A$. The position of the vectors $C_j$ in $P$ is identical to the position of the associated eigenvalue on the diagonal of $D$. This identity implies that $A$ is similar to $D$. Therefore, $A$ is diagonalizable.

   **Remark 2.4.** If the algebraic multiplicity $n_i$ of the eigenvalue $\lambda_i$ is equal to $1$, then obviously we have $m_i = 1$. In other words, $n_i = m_i$.

5. If for some eigenvalue the algebraic multiplicity is not equal to the geometric multiplicity, then $A$ is not diagonalizable.

**Example 2.5.** Consider the matrix

$$A = \begin{pmatrix} -1 & -1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

In order to find out whether $A$ is diagonalizable, let us follow the steps described above.

1. The polynomial characteristic of $A$ is

$$p(\lambda) = \begin{vmatrix} -1-\lambda & -1 & 1 \\ 0 & -2-\lambda & 1 \\ 0 & 0 & -1-\lambda \end{vmatrix} = (-1-\lambda)^2(-2-\lambda).$$

So $-1$ is an eigenvalue with multiplicity $2$ and $-2$ with multiplicity $1$.

2. In order to find out whether $A$ is diagonalizable, we only concentrate our attention on the eigenvalue $-1$. Indeed, the eigenvectors associated to $-1$, are given by the system

$$\Big(A + I_n\Big)X = \begin{pmatrix} 0 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} X = \mathcal{O}.$$

This system reduces to the equation $-y + z = 0$. Set $x = \alpha$ and $y = \beta$, then we have

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

So the geometric multiplicity of $-1$ is $2$ the same as its algebraic multiplicity. Therefore, the matrix $A$ is diagonalizable. In order to find the matrix $P$ we need to find an eigenvector associated to $-2$. The associated system is

$$\Big(A + 2I_n\Big)X = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} X = \mathcal{O}$$

which reduces to the system

$$\begin{cases} x - y &= 0 \\ z &= 0 \end{cases}$$

Set $x = \alpha$, then we have

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Set

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then

$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

But if we set

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

then

$$P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

We have seen that if $A$ and $B$ are similar, then $A^n$ can be expressed easily in terms of $B^n$. Indeed, if we have $A = P^{-1}BP$, then we have $A^n = P^{-1}B^nP$. In particular, if $D$ is a diagonal matrix, $D^n$ is easy to evaluate. This is one application of the diagonalization. In fact, the above procedure may be used to find the square root and cubic root of a matrix. Indeed, consider the matrix above

$$A = A = \begin{pmatrix} -1 & -1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Set

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

then

$$P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D.$$

Hence $A = PDP^{-1}$. Set

$$B = P \begin{pmatrix} -2^{1/3} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{-1},$$

Then we have

$$B^3 = A.$$

In other words, $B$ is a cubic root of $A$.